

一个基于代数不变量的目标识别算法

易英辉

宋克欧

(哈尔滨工程大学水声工程系, 哈尔滨 150001) (哈尔滨工程大学计算机与信息科学系, 哈尔滨 150001)

摘要 首先给出了用于代数不变量的“参照系”和“坐标”的概念, 然后给出了一个基于代数不变量进行目标识别的算法, 最后给出了算法的实验结果。

关键词 不变量, 参照系, 训练, 识别

0 引言

不变量理论大量引入计算机视觉研究发生于80年代末期, 直至1991年, 世界上才第一次召开这方面的研讨会。目前, 国外学者对这一课题已表现出极大兴趣和关心。

视觉不变量理论的优点和应用在于它既不介意目标平移、缩放所造成的形状相似变化, 也不介意目标滚动造成的投影形态变化。同时它可以冗余地表示目标全局结构特征, 当发生部分遮挡时, 往往可以从剩余部分测出表示全局的结构不变信息。

视觉不变量理论认为: 几何特征是机器视觉的基础, 不变量是几何结构的本质描述。所谓视觉不变量就是目标几何结构在某些变换群下保持不变的函数形式。本文给出的是一个基于不变量的识别算法。所谓“识别”, 意味着对象(识别目标)是事先已知的, 或已知它的特性, 或已知它的特征, 或已知它的形象, 或已知它的结构等。用计算机视觉、模式识别的术语, 即称这种识别是“基于模板的”(model-based)。基于模板的识别问题, 包含“建立模板库”和“识别”两个过程。

1 算法原理

在本算法限定下的目标的射影图, 是一个平面

的点的集合。如果我们能够得到集合上的一些特殊点, 并用这些特殊点提取出代数不变量, 问题就解决了。

我们知道, 一个点的自然描述就是在给定参照系下这个点的“坐标”。知道了坐标就可以得到相关的代数函数, 从而得到相关的代数不变量。

在本识别算法中, 所选用的“参照系”和“坐标”必须满足以下要求:

(1) 映射不变性。即在射影变换下“参照系”和“坐标”必须保持不变, 这样才能处理不同视角下目标的识别。

(2) 模板和待识别图象必须选择同一参照系, 这样才能解决匹配的问题。

对于任意给定的5个点 p_1, \dots, p_5 我们可以得到2个不变量:

$$I_1(p_1, p_2, p_3, p_4, p_5) = \frac{|m_{431}| |m_{521}|}{|m_{421}| |m_{531}|}$$
$$I_2(p_1, p_2, p_3, p_4, p_5) = \frac{|m_{421}| |m_{532}|}{|m_{432}| |m_{521}|}$$

其中 $m_{ijk} = (p_i \ p_j \ p_k)$, $p_i = (x_i \ y_i \ 1)^T$, $|m|$ 为矩阵 m 的行列式(如图1)。该不变量在射影变换下保持不变。

这样, 我们选取 p_1, p_2, p_3, p_4 4个点构成的有序组作为参照系, 不变量 I_1 和 I_2 就可以看作点 p_5 在这个参照系下的“坐标”。对于任意一个属于同一个刚体的有 n 个元素的点集 ($n \geq 5$), 都可以任意选

* 国防科委基金和黑龙江省自然科学基金项目
收稿日期: 1997-12-11; 收到修改稿日期: 1998-06-01

出其中4个点构成有序组,以这个有序组作为参照系,计算出其余 $n-4$ 个点在这个参照系下的“坐标”,即计算 $I_1(p_1, p_2, p_3, p_4, p_5), I_2(p_1, p_2, p_3, p_4, p_5), I_1(p_1, p_2, p_3, p_4, p_6), I_2(p_1, p_2, p_3, p_4, p_6), \dots, I_1(p_1, p_2, p_3, p_4, p_n), I_2(p_1, p_2, p_3, p_4, p_n)$ 。

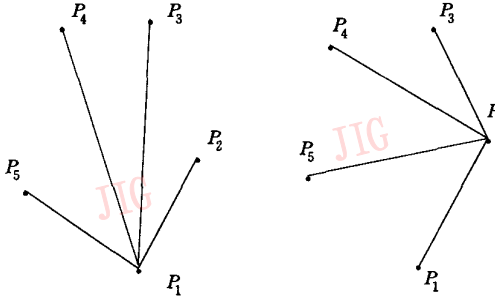


图1 平面上五个点有两个独立的不变量

上面给出的点的描述在射影变换下是不变的,即是不变量。

这样一来,对于一个有 n 个特殊点的目标,在数据库中,有 p_n^4 个参照系,对每个参照系,都计算其余 $n-4$ 个点在此参照系下的“坐标”不变量,因此共得到 p_n^5 个“坐标”不变量。对于有 m 个模板的数据库,设第 i 个模板有 n_i 个特殊点,则数据库中的“坐标”不变量为:

$$\begin{matrix} p_{11} & p_{12} & \dots & p_{1p_{n_1}^5} \\ p_{21} & p_{22} & \dots & p_{2p_{n_2}^5} \\ \vdots & \vdots & \ddots & \vdots \\ p_{m1} & p_{m2} & \dots & p_{mp_{n_m}^5} \end{matrix}$$

其中 $p_{ij}=(I_{1ij} \ I_{2ij})$ 为第 i 个模板的第 j 个“坐标”不变量。

2 模板数据库建立算法

根据上面给出的参照系——“坐标”描述点的原理,我们可得如下建立模板数据库的算法:

(1)对模板图象进行预处理,抽取模板的所有可得到的特征点(假定找到 n 个特征点)。

(2)对每一个由4个点组成的有序组(其中任何三点不共线),即对 P_n^4 个有序组,都进行如下步骤:

(2.1)以这个有序组作为映射参照系,计算其余 $n-4$ 个点在这个参照系下的“坐标”。

(2.2)用这个“坐标”值作为哈希表(hash table)的地址或索引函数,并在这个表的适当入口处,记录下这个“坐标”所基于的映射参照系(按顺序地)和这

个点所在的模板。

哈希表的建立过程可以看作是对模板的学习或训练过程,即模板库的建立过程。在这一阶段,模板的相关信息以不变量的描述方式存储起来。因而,哈希表可以看作模板目标在不同参照系下的记忆。

3 识别算法

识别阶段,要应用在学习阶段已准备好的哈希表。这样,给定一幅图象,首先抽取所有特征点,任选其中4个构成的有序组作为参照系,计算其余各点在参照系下的“坐标”,并试图将图象中点的“坐标”同哈希表中所存储的模板点的“坐标”相匹配。

下面是能处理部分遮挡目标的识别算法:

(1)对目标图象进行预处理,抽取目标的所有能得到的特征点(假定能找到 n 个点)。

(2)任选其中4个点构成有序组(其中任何三点不共线),以这个有序组作为一个参照系,计算其余 $n-4$ 个点在这个参照系下的“坐标”。

(3)对每一个这样的“坐标”,根据索引函数,直接找到其在哈希表中相对应一条记录,并得到记录中所记载的参照系和所属模板。对比模板投一票,并记录下模板中五个点的坐标(笛卡尔坐标,而非参照系下的“坐标”)与目标图象中点的坐标的对应。

(4)找到得票数最多的一个模板,此模板就是该目标识别的结果。

(5)根据模板与目标间的点的对应关系,计算出均方误差最小的两幅图象间的射影变换的参数。

(6)根据上一步所得到的射影变换,将整个模板图象作此变换,并用变换图象同目标图象进行匹配,在误差允许的范围验证识别的正确性。若验证失败,转到第2步,选取另外4点作为参照系,重新进行识别的全过程。

整个算法流程见图2。

4 算法复杂度

在整个建库步骤中,对一个模板,设有 n 个特征点,共需要计算 P_n^5 个不变量,故其计算复杂度为 $O(n^5)$ 。

在识别过程中,目标图象有 n 个特征点,每选择一个参照系,整个投票过程的计算复杂度为 $O(n)$ 。这样,总的识别时间只取决于 n ,而与数据库中的模

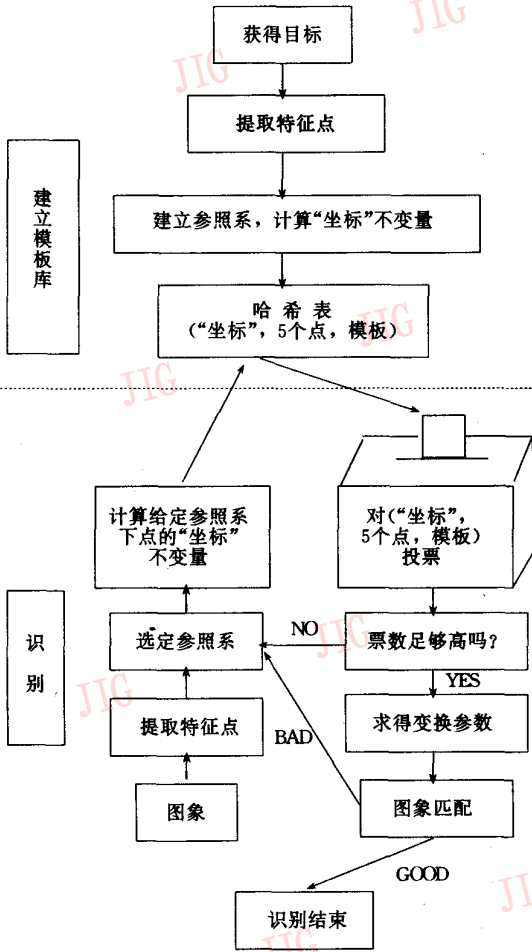


图 2 算法流程图

板数是无关系的。尽管在最坏的情况下,我们可能遍取所有可能的参照系,尤其是目标并不属于某一模板时,整个识别的计算复杂度为 $O(n^5)$,仍与模板数无关。而在大多数情况下,尤其是目标是无遮挡的,并确实属于某一模板,识别总能在第一遍的投票过程中即告成功。

5 实验结果

5.1 误差实验

由前面几节的知识我们可知,理论上不变量在各种变换下保持不变,但在实际识别时,由于噪声的干扰和图象采集分辨率的不同,不变量有少许误差。这个误差到底有多大?这直接关系到识别的正确与否,因此我们作了下面的实验:

首先选择一个五边形在扫描仪上获取的图象作为模板(如图 3a 所示),按照特征点抽取的顺序,共

可得 120 对不变量。然后利用摄像机,在不同的角度摄取这个模板的图象,并按照同样的特征点抽取顺序,计算出相应的 120 对不变量。将每幅图象计算出的不变量同模板不变量相比较,即可得出不变量在该视角下的相对误差。

本实验共取了 3 种视角,即直视(0°)(图 3c)、垂直方向 50° 仰视(图 3b)、水平方向 60° 斜视(图 3d)。

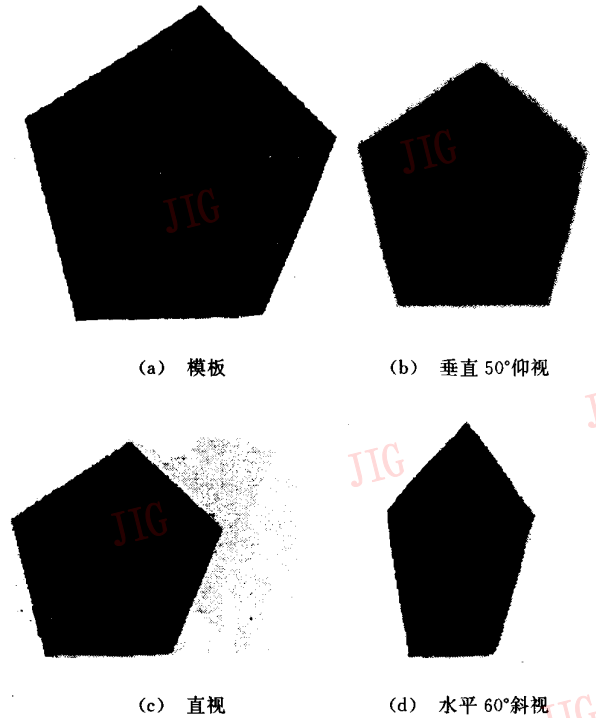


图 3 实验图象

表 1 为实验的结果分析。本表计算出的误差为不同视角下图象同模板的相对误差。

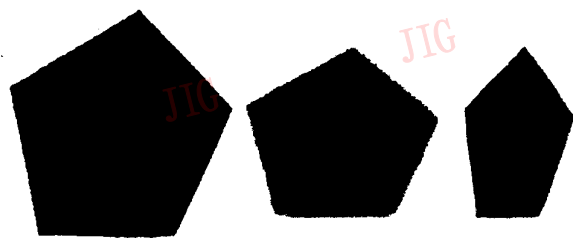
本实验表明,在不同视角下,不变量的平均误差小于 4%,这个结果可以用于识别时的投票步骤。

表 1 不变量的误差

	最大误差	最小误差	平均误差	识别结果
图 3b	5.5%	0.1%	2.3%	✓
图 3c	2.1%	0.1%	0.9%	✓
图 3d	9.0%	0.9%	3.7%	✓

5.2 识别实验

本实验共选择了四个较简单的多边形对象,对每个对象都进行了垂直 45° 仰视和水平 45° 斜视两个视角的识别。图 4、5、6、7 分别为对象一、二、三、四的模板图、仰视图和斜视图,识别结果见表 2。



(a)模板图象 (b)45°仰视 (c)45°斜视

图4 对象1



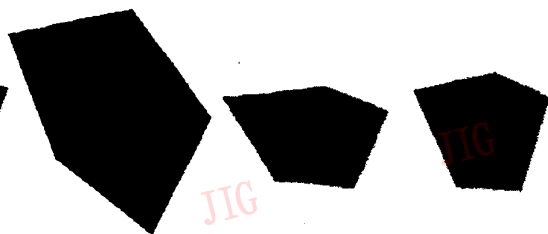
(a)模板图象 (b)45°仰视 (c)45°斜视

图5 对象2



(a)模板图象 (b)45°仰视 (c)45°斜视

图6 对象3



(a)模板图象 (b)45°仰视 (c)45°斜视

图7 对象4

表2 识别实验的结果

对象	对象1	对象2	对象3	对象4
图象	图4b 图4c	图5b 图5c	图6b 图6c	图7b 图7c
识别结果	✓	✓	✓	✓

本实验表明,运用不变量可以在不同视角对图象进行正确识别。

6 结束语

建库步骤的运行,不需要对待识别图象有任何先验知识。因而,它是单独运行的,其运行时间不被加到实际的识别时间当中去。当有新的模板加到数据库当中时,对其处理是独立的,无需重新计算整个哈希表。

在建库过程中,每个模板都要计算 P_n^5 个不变量,这样作虽然计算量很大,但有以下几点好处:

(1) 确保识别时所选参照系同建库时所选参照系相同。

(2) 以建库的时间和数据库长度为代价,来降低识别所用的时间。

(3) 便于对应点的匹配,以确定模板与待识别图象间映射变换的参数。

(4) 能够进行被部分遮挡目标的识别。

从上面的说明中我们还可以看出不变量识别应

用中的一个重要特点,即能进行部分被遮挡目标的识别。这是因为待识别图象中的一个被部分遮挡目标的点集的“坐标”,同存储的模板“坐标”有部分重合。若这些“坐标”是基于同样4个有序点组成的参照系,那么,只要目标有不少于5个点没有被遮挡,我们都可以从中选出4个点作为参照系,计算出其余点的“坐标”同模板库中的相匹配,来完成部分遮挡目标的识别。

参考文献

- 1 宋克欧,黄凤岗,易英辉. 视觉不变量与机器视觉研究. 中国图象图形学报,1997,2(7):468~470.
- 2 Mundy J L, Zisserman A. Geometric Invariance in Computer Vision. London, MIT Press, Cambridge, Ma, 1992.
- 3 Forsyth D, Mundy J L, Zisserman A, et al. Invariant descriptors for 3D object recognition and pose. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1991,13(10):971~991.
- 4 Reiss T H, Object Recognition Using Algebraic and Differential Invariants. Signal Processing, 1993,32:367~395.
- 5 Barrett E B, Payton P M, Haag N N, et al. Invariants Under Image Perspective Transformations; Theory and Examples. Int. J. Pattern Recognition Image Processing, 1990,12:296~314.
- 6 Forsyth D A, Mundy J L, Zisserman A P, et al. Invariance-A New framework for Vision. Proc. ICCV3,1990,598~605.
- 7 Lamdan Y, Schwarts J T, Wolfson H J. Object Recognition by Affine Invariant Matching. Proc. CVPR,1988,335~344.



易英辉 哈尔滨工程大学水声工程系博士研究生。主要研究领域包括:水声信号处理、图象处理、计算机视觉等。



宋克欧 哈尔滨工程大学计算机与信息科学系教授。主要研究领域包括图象处理、模式识别、计算机视觉、水声信号特征提取和识别等。

An Object Recognition Algorithm Based on Algebraic Invariants

Yi Yenghui

(Dept. of underwater Acoustic Engineering Harbin Engineering Univ. Harbin 150001)

Song Keou

(Dept. of Computer & Information Science, Harbin Engineering Univ. Harbin 150001)

Abstract Invariants can be divided into algebraic invariants and deferential invariants. This thesis gives out an algorithm to recognize objects based on algebraic invariants. It has two parts: principle and algorithm. In the principle part, the concept of 'bases' and 'coordinate' is given under recognition using algebraic invariants. The algorithm part includes 'training' and 'recognition'. Atlast the thesis gives out the experimental results of this algorithm.

Keywords Invariant, Bases, Training, Recognition

VTEL 中标全国烟草行业卫星通信网会议电视项目

全球网络化数字视频通信领域的领导者——VTEL 美国视讯公司宣布赢得了“全国烟草行业卫星通信网会议电视项目。”

VTEL 公司将向国家烟草专卖局全国卫星会议电视网提供 35 套 TC2000 多媒体数字视讯终端设备。网络建成后将覆盖全国 35 个省、市、自治区。在全网召开电视会议的同时,更可方便地进行异地远程文件共享,数据传输,电子白板等多媒体数据业务。在 VTEL 的全部产品升级适合 H. 323 标准后,更扩大了应用范围,可以轻松地接入局域网,即在支持 ITU 的 H. 320 标准的同时又支持 H. 323 标准。

此项目由中机招标公司按国际惯例面向全球范围招标。在竞标过程中,专家组对各个厂家的投标书及设备进行了近两个月的严格评审。在比较了各会议电视厂家产品的图像质量、大网组网能力、技术先进性、稳定及可靠性、升级及扩容能力、售后服务支持、价格等诸项指标之后,最终选择了 VTEL 公司代表世界会议电视技术发展趋势的智能多媒体会议电视系统 TC2000。同时,VTEL 的中标与其在中国广大用户中良好的信誉、丰富的组网经验、严谨的工作态度以及完美的售后服务是分不开的。